

Structuration de la diversité génétique chez la luzerne cultivée, conséquence pour l'identification de gènes liés à des caractères agronomiques

Julier B.¹, Herrmann D.¹, Flajoulot S.², Barre P.¹, Huyghe C.¹, Ronfort J.³

¹ INRA, UR4, Unité de Recherche Pluridisciplinaire Prairies et Plantes Fourragères, 86600 Lusignan, ² GIE Grass, La Litière, 86600 Saint-Sauvant, ³ INRA, UMR Amélioration Génétique et Adaptation des Plantes méditerranéennes et Tropicales Montpellier 2 place Viala 34060 Montpellier

Correspondance : Bernadette.Julier@lusignan.inra.fr

Résumé

La luzerne, espèce allogame et autotétraploïde, est une légumineuse fourragère dont l'intérêt connaît un renouveau certain grâce à ses atouts pour le développement durable de l'agriculture. L'amélioration des variétés nécessite de mieux connaître les ressources génétiques, pour les exploiter, que ce soit en sélection classique ou en utilisant les outils moléculaires pour implémenter la sélection assistée par marqueurs. En étudiant 10 variétés européennes, nous avons montré que la diversité à l'intérieur des variétés est grande (hétérozygotie attendue de 0.75, déviation standard de 0.80, pour des caractères phénotypiques et des marqueurs moléculaires neutres, respectivement). La différenciation entre variétés est notable pour les caractères phénotypiques. Cette structuration de la diversité est favorable à l'utilisation de la génétique d'association basée sur des gènes candidats pour identifier des gènes liés à des caractères agronomiques. Une telle étude a permis de montrer que le gène *Constans*-like contribue à expliquer les différences de longueur de tiges chez la luzerne, une composante du rendement fourrager.

Mots-clés : légumineuse, marqueur moléculaire, phénotype, variété, population de pays.

Abstract:

Alfalfa, an allogamous and autotetraploid species, is a forage legume for which a renewed interest has raised because of its advantages for a sustainable agriculture. Variety improvement requires a better knowledge of the genetic resources in order to exploit them, with classical breeding methods or with molecular tools to implement marker assisted selection. We have described 10 European varieties and found a large within-variety diversity (expected heterozygosity of 0.75, standard deviation of 0.80, with phenotypic traits and neutral molecular markers, respectively). Differentiation between varieties was noticeable for phenotypic traits. This genetic structure of diversity is favourable to the use of association genetics based on candidate genes to identify genes related to agronomic traits. Such a study showed that *Constans*-like gene contributes to explain differences in alfalfa stem length, a component of forage yield.

Keywords: legume, molecular marker, phenotype, variety, landrace

1- Introduction

Mieux connaître la diversité génétique et la structuration des populations chez la luzerne cultivée est utile pour la gestion des ressources génétiques et pour l'amélioration des variétés, que ce soit par les méthodes traditionnelles ou en utilisant des outils moléculaires. La sélection assistée par marqueurs (Meuwissen et

al., 2001) ou la sélection génomique (Hayes et al., 2013) nécessitent de travailler avec du matériel génétique contenant des sources de progrès génétique et adapté aux méthodes mises en œuvre (populations biparentales de cartographie, populations à bases plus ou moins larges pour la génétique d'association ou la sélection génomique).

L'objectif de cette étude était de caractériser la diversité des deux types de variétés de luzerne (Flamande et Provence) avec des marqueurs moléculaires et des données morphologiques. Les résultats obtenus sont analysés dans la perspective d'identifier des gènes liés aux caractères agronomiques.

2- Matériel et méthodes

Dix variétés européennes ont été étudiées : 3 variétés (dont une population de pays) de type Provence ayant une faible dormance et issues de la sous-espèce *M. sativa sativa*, et 7 variétés (dont une population de pays) de type Flamande, un pool plus dormant, introgressé par la sous-espèce *M. sativa falcata*. Pour chaque variété, 40 plantes, correspondant à 40 génotypes ont été échantillonnées. Après bouturage, elles ont été étudiées en deux lieux, avec 3 répétitions par lieu en pépinière de plantes isolées pour 4 caractères mesurés à plusieurs coupes au cours de 4 années : date de floraison, port, vitesse d'allongement des tiges, longueur maximale des tiges. Une analyse canonique discriminante a été effectuée, et des distances de Mahalanobis ont été calculées entre variétés. Un indice de diversité intra-variétale a été estimé en utilisant la déviation standard pondérée des 5 premiers axes de l'analyse canonique. Les génotypes ont aussi été analysés avec 8 marqueurs SSR (FMT13, MTIC451, MTIC338, MAA660456, B14B03, MTIC343, MTIC432, MTIC82 cartographiés sur les chromosomes 1, 2, 3, 4, 5, 6, 7 et 7, respectivement) (Juier et al., 2003) dont la dose des allèles a été déterminée. La diversité intra-variétale a été évaluée par l'hétérozygotie attendue. La différenciation entre variétés a été testée par le F_{ST} entre paires de variétés. La pertinence d'analyser 40 génotypes par variété a été analysée. Pour cela, sur 2 variétés (Mercedes et Symphonie), 120 individus par variété ont été échantillonnés. Ils ont été génotypés avec deux marqueurs microsatellites (MAA660456 et MTIC432). Vingt sous-échantillons de 10, 20 et 40 individus ont été aléatoirement tirés. Le nombre d'allèles obtenus dans chaque tirage a été évalué. Les tirages ont été utilisés pour calculer des F_{ST} par paire de sous-échantillons, de la même variété ou de variétés différentes.

3- Résultats et discussion

Echantillonnage des variétés

Plus la taille d'échantillon est petite, plus le nombre total d'allèles identifiés est faible (Tableau 1). Parmi ces allèles, certains sont rares (fréquence supérieure à 1%), et il est généralement admis qu'ils biaisent les analyses de diversité. Quand on considère les allèles non rares, on constate que l'échantillon de 40 individus contient presque autant d'allèles que le lot de 120 individus, alors que l'échantillon de 10 individus induit la perte d'allèles non rares. L'échantillon de 20 individus occasionne quelques pertes d'allèles. Les valeurs de F_{ST} entre variétés étaient toujours significatives quand des échantillons de 40 individus étaient utilisés. Par contre, avec des échantillons de 10 individus, plus de la moitié des F_{ST} étaient non significatifs. Avec des échantillons de 20 individus, les F_{ST} étaient significatifs dans 18% des cas avec le marqueur MAA660456 et dans 82% des cas avec le marqueur MTIC432 (Herrmann et al. 2010a). Nous avons conclu de cette étude qu'un échantillon de 40 individus était de taille raisonnable pour des études de diversité entre des variétés de luzerne tétraploïde relativement semblables. Si des contraintes techniques ou économiques réduisent les capacités d'analyse, des échantillons d'une taille de 20 individus peuvent être corrects, mais des échantillons de 10 individus ne permettent pas d'obtenir des résultats fiables.

Tableau 1 : Pour deux marqueurs SSR (MAA660456 and MTIC432) et deux variétés (Mercedes and Symphonie), nombre d'allèles observés chez 120 génotypes et nombre moyen d'allèles dans 20 sous-échantillons aléatoires de différentes tailles (40, 20, 10 génotypes). Le nombre d'allèles est soit un nombre total, soit le nombre d'allèles dont la fréquence est supérieure à 1%.

	Mercedes		Symphonie	
	MAA660456	MTIC432	MAA660456	MTIC432
<i>Avec les 120 génotypes</i>				
Nombre d'allèles	12	16	9	17
Nombre d'allèles >1%	7	9	8	9
<i>Sous-échantillons de 40 génotypes</i>				
Nombre moyen d'allèles	9.9	12.5	8.5	13.8
Nombre d'allèles >1%	6.95	8.95	7.90	8.95
<i>Sous-échantillons de 20 génotypes</i>				
Nombre moyen d'allèles	8.6	10.5	7.9	10.9
Nombre d'allèles >1%	6.85	8.55	7.55	8.25
<i>Sous-échantillons de 10 génotypes</i>				
Nombre moyen d'allèles	7.0	8.7	7.0	8.5
Nombre d'allèles >1%	6.15	7.55	6.85	7.00

Diversité à l'intérieur des variétés

Les caractères morphologiques et les marqueurs moléculaires révèlent une large diversité intra-variétale (Tableau 2). Les 8 variétés et les 2 populations de pays contiennent le même niveau de diversité intra-variétale, ce qui signifie que la création variétale n'a pas restreint la diversité présente dans une variété. De façon similaire, la diversité des variétés issues d'un programme de sélection d'un obtenteur privé est élevée et n'a pas été réduite au cours des cycles de sélection (Flajoulot et al. 2005). Cette diversité provient du mode d'obtention des variétés synthétiques, dans lesquels de nombreux parents sont utilisés dans le polycross initial, ces parents étant en général eux-mêmes des familles de demi-frères.

Tableau 2 : Indices décrivant la diversité intra-variétale pour les 10 variétés: hétérozygotie attendue (H_E) avec 8 marqueurs SSR et déviation standard pondérée des 5 premiers axes d'une analyse canonique effectuée sur les caractères morphologiques (DS).

Variété	Type	H_E	DS
Flamande	Flamande	0.751	0.81
Luzelle	Flamande	0.760	0.93
Mercedes	Flamande	0.736	0.84
Alpha	Flamande	0.761	0.76
Symphonie	Flamande	0.762	0.85
Cannelle	Flamande	0.726	0.84
Harpe	Flamande	0.740	0.90
Provence	Provence	0.750	0.88
Zénith	Provence	0.762	0.84
Barmed	Provence	0.745	0.81

Dans cette configuration, la diversité neutre est élevée. Deux hypothèses peuvent expliquer l'importance de la diversité phénotypique (non neutre) : (1) le fardeau génétique généré par des allèles délétères est important chez cette espèce allogame et autotétraploïde, ce qui empêche les sélectionneurs de pratiquer

des générations de consanguinité pour fixer les allèles et donc les caractères favorables, (2) cette variabilité phénotypique a un rôle fonctionnel dans la valeur agronomique des variétés, la combinaison au sein d'une même population d'individus ayant des caractéristiques sensiblement différentes permettant une meilleure adaptation aux contraintes du milieu.

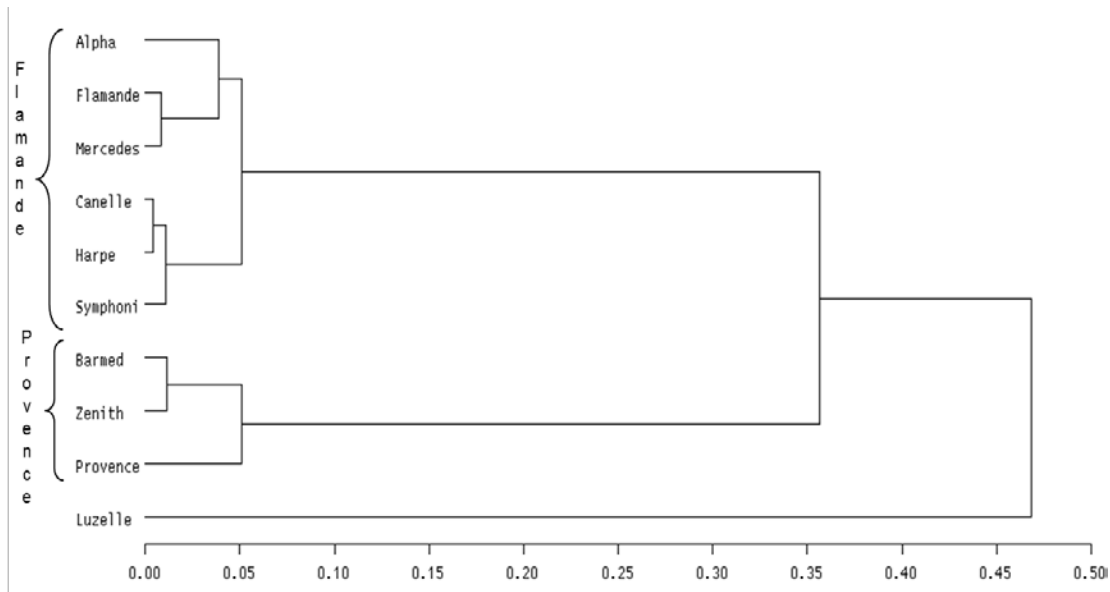
Diversité entre variétés

Les caractères morphologiques révèlent des différences faibles, bien que significatives entre variétés (Tableau 3). Avec les caractères morphologiques, la séparation est nette entre variétés Flamande et Provence. Luzelle, une variété qui contient des parents issus de la sous-espèce *falcata*, est bien séparée des autres variétés (Tableau 3, Figure 1). Entre les variétés de type Flamande, la différenciation est inexistante, alors que les trois variétés de type Provence sont contrastées pour les caractères morphologiques. Avec les marqueurs moléculaires, on a trouvé une différenciation entre Luzelle et les autres variétés, comme avec les caractères morphologiques. Barmed, une variété non dormante sélectionnée dans du matériel américain et nord-africain, est différente des autres variétés. Aucune structuration entre les types Flamande et Provence n'a été détectée avec les marqueurs moléculaires (Tableau 3, Figure 2). Finalement, même si la structuration entre variétés est différente avec les caractères morphologiques et les marqueurs moléculaires, une grande diversité intra-variété a toujours été observée et Luzelle est différente des autres variétés.

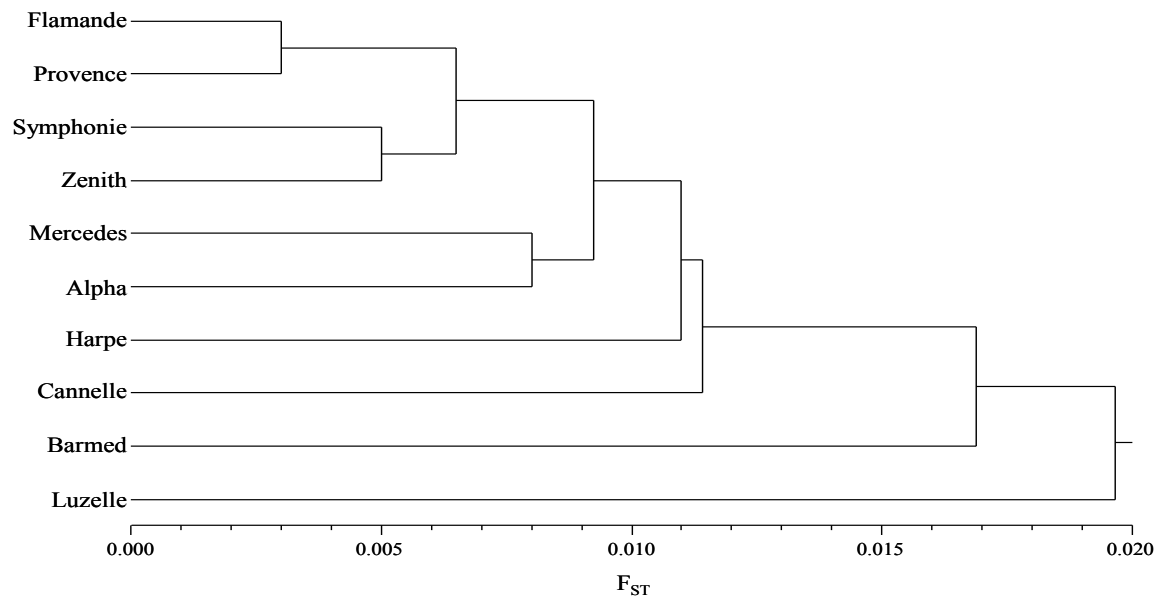
Ces résultats montrent que la structuration des variétés de luzerne à l'aide de quelques marqueurs moléculaires neutres n'est pas efficace. Il se peut, et cela nécessite de nouvelles études, que l'utilisation d'un grand nombre de marqueurs, issus de techniques SNP (Han et al., 2011; Li et al., 2014) ou de génotypage par séquençage (Elshire et al., 2011) permettent une meilleure structuration du matériel génétique. Cependant, la faible structuration intra-variétale ou inter-variétale pour les marqueurs neutres est favorable à la mise en œuvre d'études de **génétique d'association** qui consistent à relier le polymorphisme de l'ADN à un polymorphisme phénotypique. Le déséquilibre de liaison, faible sur une espèce allogame comme la luzerne, indique qu'une **approche « gène candidat »** doit être privilégiée.

Tableau 3 : Distances entre les 10 variétés. Au-dessus de la diagonale: distances de Mahalanobis calculées avec les caractères morphologiques. En dessous de la diagonale : valeurs F_{ST} calculées avec les marqueurs SSR. F_{ST} et distances de Mahalanobis sont significatives à $P < 0.01$ sauf la paire Harpe/Cannelle pour la distance de Mahalanobis qui est significative à $P < 0.05$.

	Flamande	Luzelle	Mercedes	Alpha	Symphonie	Cannelle	Harpe	Provence	Zénith	Barmed
Flamande	*****	16.37	4.56	6.62	7.94	6.79	7.96	8.47	10.40	9.54
Luzelle	0.020	*****	13.41	21.66	17.01	16.11	18.99	18.66	24.74	27.24
Mercedes	0.006	0.026	*****	5.55	5.32	7.45	6.55	7.78	8.43	9.96
Alpha	0.008	0.019	0.008	*****	6.82	6.93	7.72	13.02	10.62	11.81
Symphonie	0.007	0.021	0.013	0.010	*****	5.46	4.33	13.82	13.32	14.54
Cannelle	0.009	0.018	0.013	0.012	0.013	*****	3.59	12.92	16.92	15.53
Harpe	0.008	0.020	0.013	0.008	0.015	0.012	*****	11.43	15.20	12.66
Provence	0.003	0.017	0.009	0.010	0.009	0.011	0.007	*****	9.21	8.49
Zénith	0.004	0.017	0.009	0.009	0.005	0.010	0.015	0.006	*****	7.28
Barmed	0.016	0.019	0.015	0.014	0.015	0.028	0.019	0.015	0.013	*****

Figure 1 : Dendrogramme établi avec les distances de Mahalanobis calculées sur les caractères phénotypiques.

Une telle étude a été menée sur ce matériel en utilisant le gène Constans-like identifié chez l'espèce modèle *Medicago truncatula* pour expliquer des différences de date de floraison et de longueur de tiges (Pierre et al., 2011). Une étude d'association entre le polymorphisme phénotypique pour la longueur des tiges et le polymorphisme de séquence du gène Constans-like chez la luzerne a été réalisée sur la population de 400 génotypes décrite ci-dessus. Nous avons montré que Constans-like explique aussi des différences pour ces caractères chez la luzerne (Herrmann et al., 2010b). Une sélection divergente pour la présence ou l'absence de l'allèle favorable à un locus permet de créer une différence de longueur de tiges de 10 à 15 cm entre les populations (Julier 2012). Ces résultats permettent d'envisager le développement de marqueurs liés à ce gène pour une sélection assistée par marqueurs.

Figure 2 : Dendrogramme établi avec les F_{ST} calculés sur les marqueurs moléculaires (SSR).

L'histoire de l'introduction de la luzerne en Europe (17^{ème} siècle), le mode de reproduction de cette espèce (allogamie), ses caractéristiques génétiques (autotétraploïdie) et les échanges traditionnels de graines entre agriculteurs et maintenant entre sélectionneurs contribuent à expliquer la faible différenciation des variétés. Cette caractéristique est finalement plutôt favorable à l'implémentation de la génétique d'association ou de la sélection génomique.

Références bibliographiques

- Elshire R.J., Glaubitz J.C., Sun Q., Poland J.A., Kawamoto K., Buckler E.S., Mitchell S.E., 2011. A robust, simple Genotyping-by-Sequencing (GBS) approach for high diversity species. *Plos One* 6.
- Flajoulot S., Ronfort J., Baudouin P., Barre P., Huguet T., Huyghe C., Julier B., 2005. Genetic diversity among alfalfa (*Medicago sativa*) cultivars coming from a single breeding program, using SSR markers. *Theoretical and Applied Genetics* 111, 1420-1429.
- Han Y.H., Kang Y., Torres-Jerez I., Cheung F., Town C.D., Zhao P.X., Udvardi M.K., Monteros M.J., 2011. Genome-wide SNP discovery in tetraploid alfalfa using 454 sequencing and high resolution melting analysis. *BMC Genomics* 12.
- Hayes B.J., Cogan N.O.I., Pembleton L.W., Goddard M.E., Wang J.P., Spangenberg G.C., Forster J.W., 2013. Prospects for genomic selection in forage plant species. *Plant Breeding* 132,133-143.
- Herrmann D., Flajoulot S., Julier B., 2010a. Sample size for diversity studies in tetraploid alfalfa (*Medicago sativa*) based on codominantly coded SSR markers. *Euphytica* 171, 441-446.
- Herrmann D., Barre P., Santoni S., Julier B., 2010b. Association of a CONSTANS-LIKE gene to flowering and height in autotetraploid alfalfa. *Theoretical and Applied Genetics* 121, 865-876.
- Julier B., Flajoulot S., Barre P., Cardinet G., Santoni S., Huguet T., Huyghe C., 2003. Construction of two genetic linkage maps in cultivated tetraploid alfalfa (*Medicago sativa*) using microsatellite and AFLP markers. *BMC Plant Biology* 3:9.
- Julier B., 2012. Alfalfa breeding benefits from genetic analyses on *M. truncatula*. In: *Proceedings of the 7th International Symposium on the Molecular Breeding of Forage and Turf*. Salt Lake City, Etats-Unis. Washington, USA: USDA-ARS, p 17-19.
- Li X.H., Han Y.H., Wei Y.L., Acharya A., Farmer A.D., Ho J., Monteros M.J., Brummer E.C., 2014. Development of an alfalfa SNP array and its use to evaluate patterns of population structure and linkage disequilibrium. *Plos One* 9.
- Meuwissen T.H.E., Hayes B.J., Goddard M.E., 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819-1829.
- Pierre J.B., Bogard M., Herrmann D., Huyghe C., Julier B., 2011. A CONSTANS-like gene candidate that could explain most of the genetic variation for flowering date in *Medicago truncatula*. *Molecular Breeding* 28, 25-35.